

# Interacting Conceptual Spaces I :

## Grammatical Composition of Concepts \*

Joe Bolt   Bob Coecke   Fabrizio Genovese   Martha Lewis   Dan Marsden   Robin Piedeleu †

University of Oxford

March 27, 2017

### Abstract

The categorical compositional approach to meaning has been successfully applied in natural language processing, outperforming other models in mainstream empirical language processing tasks. We show how this approach can be generalized to conceptual space models of cognition. In order to do this, first we introduce the category of convex relations as a new setting for categorical compositional semantics, emphasizing the convex structure important to conceptual space applications. We then show how to construct conceptual spaces for various types such as nouns, adjectives and verbs. Finally we show by means of examples how concepts can be systematically combined to establish the meanings of composite phrases from the meanings of their constituent parts. This provides the mathematical underpinnings of a new compositional approach to cognition.

## 1 Introduction

How should we represent concepts and how can they be composed to form new concepts, phrases and sentences? These questions are fundamental to cognitive science. *Conceptual spaces theory* gives a way of describing structured concepts [12, 13], not starting from linguistic assumptions, but from cognitive considerations about human reasoning. Conceptual spaces describe a “semantics of the mind”, modelling mental descriptions of concepts. The key idea is that human beings represent concepts geometrically in certain fundamental domains of understanding such as space, motion, taste and colour. These domains are combined to form a *conceptual space* describing the features of interest. Conceptual spaces provide a middle ground between symbolic and connectionist representations of concepts, oriented towards tasks of interest in cognitive science, such as categorization and assessing similarity. A concept is then described by convex subsets of the relevant domains. The convexity requirement can be seen as a means of identifying robust, meaningful concepts. For example, if two points in some space are considered to represent the colour red, then intuitively we would expect that every point “in between” would also be considered red.

*Categorical compositional distributional models* [11] successfully exploit the compositional structure of natural language in a principled manner, and have outperformed other approaches in Natural Language Processing (NLP) [14, 18]. The approach works as follows. A mathematical formalization of grammar is chosen, for example *Lambek’s pregroup grammars* [21], although the approach is equally effective with other categorial grammars [7]. Such a categorial grammar allows one to verify whether a phrase of a sentence is grammatically well-formed by means of a computation that establishes the overall grammatical type, referred to as a *type reduction*. The meanings of *individual* words are established using a distributional model of language, where they are described as vectors of co-occurrence statistics derived automatically from corpus data [23]. The categorical compositional distributional programme

---

\*This paper is a significantly extended version of the workshop paper [3]

†authors in alphabetical order

unifies these two aspects of language in a compositional model where grammar mediates composition of meanings. This allows us to derive the meaning of sentences from their grammatical structure, and the meanings of their constituent words. The key insight that allows this approach to succeed is that both pregroups and the category of vector spaces carry the same abstract structure [11], and the same holds for other categorial grammars since they typically have a weaker categorial structure.

The abstract framework of the categorial compositional scheme is actually broader in scope than NLP. It can be applied in other settings in which we wish to compose meanings in a principled manner, guided by structure. The outline of the general programme is as follows:

1. (a) Choose a compositional structure, such as a pregroup or any other categorial grammar.  
(b) Interpret this structure as a category, the **grammar category**.
2. (a) Choose or craft appropriate meaning or concept spaces, such as spaces of propositions, vector spaces, density matrices [25, 2] or conceptual spaces.  
(b) Organize these spaces into a category, the **semantics category**, with the same abstract structure as the grammar category.
3. Interpret the compositional structure of the grammar category in the semantics category via a functor preserving the type reduction structure.
4. Bingo! This functor maps type reductions in the grammar category onto algorithms for composing meanings in the semantics category.

In order to move away from vector spaces, we construct a new categorial setting for interpreting meanings which respects the important convex structure emphasized in conceptual spaces theory. We show that this category has the necessary abstract structure required by categorial compositional models. We then construct convex spaces for interpreting the types for nouns, adjective and verbs. Finally, this allows us to use the reductions of the pregroup grammar to compose meanings in conceptual spaces. We illustrate our approach with concrete examples, and go on to discuss directions for further research.

## 2 Categorial compositional meaning

In this section we describe the details of the categorial compositional approach to meaning. We provide examples of semantic categories and grammar categories, and show the general method by which the grammar category induces a notion of concept composition in the semantic category.

### 2.1 Monoidal categories

We begin by briefly review some of the category theory underpinning categorial compositional models.

**Definition 1.** A **monoidal category** is a tuple  $(C, \otimes, I, \alpha, \lambda, \rho)$  where

- $C$  is a category
- $\otimes$ , the **tensor**, is a functor  $C \times C \rightarrow C$  where we write  $A \otimes B$  for  $\otimes(A, B)$
- $I$ , the **unit**, is an object of  $C$

The remaining data are natural isomorphisms, with components of type:

- $\alpha_{A,B,C} : ((A \otimes B) \otimes C) \rightarrow (A \otimes (B \otimes C))$
- $\rho_A : A \otimes I \rightarrow A$
- $\lambda_A : I \otimes A \rightarrow A$

These natural isomorphisms, moreover, must be such that any formal and well-typed diagram made up from  $\otimes, \alpha, \lambda, \rho, \alpha^{-1}, \rho^{-1}, \lambda^{-1}$  and identities commutes. Here ‘formal’ means ‘not dependent on the structure of any particular monoidal category.’

For a precise statement and discussion of the above definition, we direct the reader to [24]. A more gentle introduction can be found in [9]. For the purpose of our paper, the objects of a monoidal category should be thought of as *system types*. A morphism  $f : A \rightarrow B$  is then a process taking inputs of type  $A$  and giving outputs of type  $B$ . The object  $A \otimes B$  represents the systems  $A$  and  $B$  composed in *parallel*. Hence, a morphism  $f \otimes g : A \otimes B \rightarrow C \otimes D$  is to be thought of as running the process  $f : A \rightarrow C$  whilst running the process  $g : B \rightarrow D$ . The object  $I$  is thought of as the trivial system.

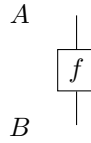
**Example 1.** The category **Rel** of sets and relations is monoidal. The tensor  $\otimes$  is the Cartesian product and  $I$  is any singleton set  $\{\star\}$ .

**Example 2.** The category **FdVect** $_{\mathbb{R}}$  of finite dimensional real vector spaces and linear maps is monoidal. The tensor  $\otimes$  is the tensor product, the trivial system  $I$  is the one-dimensional real vector space  $\mathbb{R}$ .

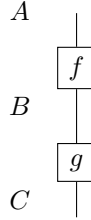
Monoidal categories admit an elegant and powerful graphical notation that we will exploit heavily in later sections. In this notation, an object  $A$  is denoted by a wire:



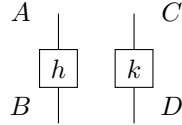
A morphism  $f : A \rightarrow B$  is represented by a box:



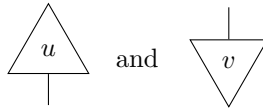
If  $g : B \rightarrow C$ , the composite  $g \circ f : A \rightarrow C$  is given by:



If  $h : A \rightarrow B$  and  $k : C \rightarrow D$ , the morphism  $h \otimes k : A \otimes C \rightarrow B \otimes D$  is depicted by:



The trivial system  $I$  is the empty diagram. Morphisms  $u : I \rightarrow A$  and  $v : A \rightarrow I$  are drawn respectively as



These special morphisms are referred to as **states** and **effects**.

## 2.2 Compact closed categories

A specific class of monoidal categories, the compact closed categories, will be of particular importance.

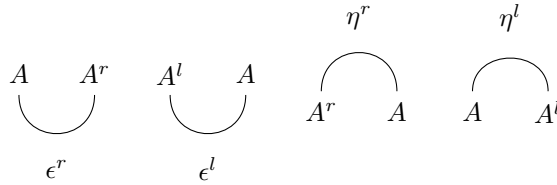
**Definition 2.** A monoidal category  $(C, \otimes, I)$  is **compact closed** if for each object  $A \in C$  there are objects  $A^l, A^r \in C$  (the **left** and **right duals** of  $A$ ) and morphisms

$$\begin{aligned} \eta_A^l : I &\rightarrow A \otimes A^l & \eta_A^r : I &\rightarrow A^r \otimes A \\ \epsilon_A^l : A^l \otimes A &\rightarrow I & \epsilon_A^r : A \otimes A^r &\rightarrow I \end{aligned}$$

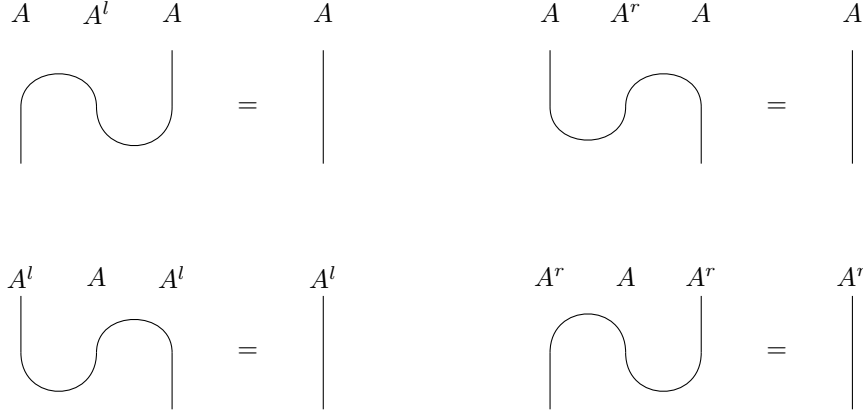
satisfying the snake equations

$$\begin{aligned} (1_A \otimes \epsilon^l) \circ (\eta^l \otimes 1_A) &= 1_A & (\epsilon^r \otimes 1_A) \circ (1_A \otimes \eta^r) &= 1_A \\ (\epsilon \otimes 1_{A^l}) \circ (1_{A^l} \otimes \eta^l) &= 1_{A^l} & (1_{A^r} \otimes \epsilon^r) \circ (\eta^r \otimes 1_{A^r}) &= 1_{A^r} \end{aligned}$$

The  $\epsilon$  and  $\eta$  maps are called **caps** and **cups** respectively, and are depicted graphically as:



Graphically, the snake equations become



Compact closed categories are a convenient level of abstraction at which to work. Many of the categories one would think to use as either grammar or meaning categories have a compact closed structure.

**Example 3.** All objects in **Rel** are self-dual. Both caps are given by

$$\epsilon_X : X \times X \rightarrow \{\star\} :: \{(x, x), \star\} \mid x \in X\}$$

The associated cup  $\eta_X$  is the converse of the above. The snake equations can be verified by direct calculation.

**Example 4.** **FHilb** is the category of finite dimensional real inner product spaces. As in the case of **FdVect** $_{\mathbb{R}}$ , the tensor  $\otimes$  is the tensor product of vector spaces and  $I$  is the one-dimensional space  $\mathbb{R}$ . In defining cups and caps, we make use of the fact that if  $\{v_i\}_i$  and  $\{u_j\}_j$  are bases for vector spaces  $V$  and  $U$  respectively, then  $\{v_i \otimes u_j\}_{i,j}$  is a basis for  $V \otimes U$ . Moreover, any linear map is fully determined by its action on a basis. Every finite-dimensional vector space is self-dual, and the cups and caps are

given by

$$\epsilon_V : V \otimes V \rightarrow \mathbb{R} :: \sum_{i,j} c_{i,j} (v_i \otimes v_j) \mapsto \sum_{i,j} c_{i,j} \langle v_i | v_j \rangle$$

$$\eta_V : \mathbb{R} \rightarrow V \otimes V :: 1 \mapsto \sum_i (v_i \otimes v_i)$$

Verifying that these maps satisfy the snake equations again follows from a straightforward calculation.

**Remark 1.** The tensor in a compact closed category is *not* necessarily a categorical product. Specifically, we cannot expect to describe every state of  $A \otimes B$  as a tensor of two states taken from  $A$  and  $B$ . We can understand this as showing composite systems have interesting behaviour that cannot be explained in terms of the behaviour of their component parts. In fact, if the tensor of a compact closed category happens to be a categorical product, then that category must be trivial, in a precise mathematical sense [8].

### 2.3 Grammar categories

Many algebraic gadgets exist to model grammar, as detailed in [6] for example. In the present work, we use Lambek's pregroup grammars, as many grammars have a pregroup structure [28]. Moreover, pregroups can be viewed as compact closed categories. It should be emphasized though that our approach does not depend on pregroups, and can be applied to other grammatical models.

**Definition 3.** A **pregroup** is a tuple  $(A, \cdot, 1, -^l, -^r, \leq)$  where  $(A, \cdot, 1, \leq)$  is a partially ordered monoid and  $-^r, -^l$  are functions  $A \rightarrow A$  such that  $\forall x \in A$ ,

$$x \cdot x^r \leq 1 \tag{1}$$

$$x^l \cdot x \leq 1 \tag{2}$$

$$1 \leq x^r \cdot x \tag{3}$$

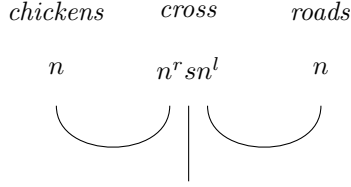
$$1 \leq x \cdot x^l \tag{4}$$

The  $\cdot$  will usually be omitted, writing  $xy$  for  $x \cdot y$ . One interprets grammar by freely generating a pregroup from a set of atomic linguistic types. Words are then assigned an element of the pregroup depending on their linguistic function. A string of words is then mapped to an element of the pregroup by multiplying together the elements associated with its constituent words in their syntactic order. If  $s_1 \cdots s_n \leq t$ , we say that the type  $s_1 \cdots s_n$  **reduces** to the type  $t$ . The pregroup freely generated by a set  $A$  is denoted by  $\text{Preg}_A$ .

**Example 5.** For simplicity, we only use the linguistic types  $n$ , for noun, and  $s$ , for sentence. Hence, we will work with  $\text{Preg}_{\{n,s\}}$ . Consider the sentence ‘*Chickens cross roads.*’ The nouns *chickens* and *roads* are of type  $n$ , and the transitive verb *cross* is assigned the type  $n^r s n^l$ . ‘*Chickens cross roads*’ therefore has type  $n(n^r s n^l)n$ . Then we have the following type reductions:

$$\begin{aligned} n(n^r s n^l)n &= (nn^r)s(n^l n) \\ &\stackrel{(1)}{\leq} s(n^l n) \\ &\stackrel{(2)}{\leq} s \end{aligned}$$

Note that we could also have performed these two steps in the opposite order. The above reduction can be given a neat graphical interpretation as follows:



where it is now very clear that the order of the reductions doesn't matter. This is a feature that is typical for pregroup grammars, while other categorial grammars such as Lambek's original categorial grammar [20] have more constraints on the order of the reductions.

A pregroup can be considered as a compact closed category. The objects of this category are the elements of the pregroup. The morphisms are given by the order structure of the pregroup. That is, there is a unique morphism  $p \rightarrow q$  if and only if  $p \leq q$ . The tensor  $\otimes$  is the monoid multiplication and the monoidal unit is the element 1. Unsurprisingly, the left and right duals of  $p$  are  $p^l$  and  $p^r$  respectively. The cups and caps are the unique morphisms given by the pregroup axioms (1)–(4). In the reduction diagram, note how the cups correspond to the cups of the compact closed structure.

## 2.4 Meaning categories

Distributional models of meaning use vector spaces to represent the meaning of words. In this paper we move away from this approach, choosing instead to work in **ConvexRel**, the category of convex sets and convexity-respecting relations. Before describing this new setting, we first present the conventional vector space approach as a demonstration of the categorical compositional method. This will prepare the ground for detailed discussion of **ConvexRel** to later sections.

One way to model meanings in a vector space is to use co-occurrence statistics [4]. The meaning of a word is identified with the frequency with which it appears near other words. One first chooses a collection of *context words*. These will be the basis vectors. One then analyses a large corpus of writing to determine how often words co-occur with the context words. For example, suppose the word *dog* appears in the same context as *cat*, *companion*, and *cuisine* respectively 19, 25, and 2 times. If our collection of context words is  $\{cat, companion, cuisine\}$ , then *dog* would be assigned the vector  $(19, 25, 2)$ . A drawback of the co-occurrence approach is that antonyms appear in similar contexts and, hence, words such as 'win' and 'lose' are indistinguishable (despite evidently being different in meaning). Another related difficulty is that vector spaces are notoriously bad for representing basic propositional logic. Nonetheless, the vector space model is highly successful in NLP.

One can also use basis vectors to represent *quality dimensions*. For example, [27, 31, 15] all represent concepts as feature vectors, with basis dimensions representing attributes of the concept. So, for instance, one might represent the word *dog* with certain values on the *fluffy*, *loyal*, and *wolf-like*. This approach closely resembles ours in this paper, though we move away from the vector space setting.

## 2.5 Putting it all together

We now have all the necessary components: a grammar category and a meaning category. We used the examples of pregroup grammars and vector spaces, but we stress yet again that the abstract method applies equally well to any two compact closed categories.

Suppose we have a string of words  $w_1 \cdots w_n$  and a pregroup  $P$ . Suppose further that  $w_i \in W_i$ , where  $W_i$  is the vector space associated with  $w_i$ 's linguistic type. The meaning of  $w_1 \cdots w_n$  is computed as follows:

1. Assign a pregroup element  $p_i$  to each word  $w_i$  based on its linguistic type.
2. Apply pregroup reduction rules (cups and caps) to the element  $p_1 p_2 \cdots p_n$  to obtain a simpler type  $x$  such that:

$$p_1 p_2 \cdots p_n \leq x.$$

3. Thinking of the above reduction as a morphism in the pregroup built up from  $\epsilon, \eta$  and identities (as given by its reduction diagram), apply the corresponding vector space morphism of type:

$$W_1 \otimes \cdots \otimes W_n \rightarrow W_X$$

to the string of word meanings represented as  $w_1 \otimes \cdots \otimes w_n$ .

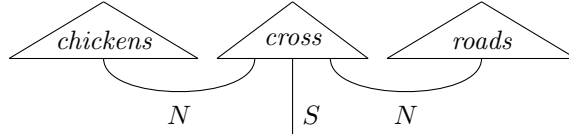
**Example 6.** Consider again the sentence ‘*chickens cross roads*’. The nouns *chickens* and *roads* have type  $n$  and so are represented in some vector space  $N$  of nouns. The transitive verb *cross* has type  $n^r sn^l$  and, hence, is represented by a vector in the vector space  $N \otimes S \otimes N$  where  $S$  is a vector space modelling sentence meaning. The meaning of ‘*chickens cross roads*’ is the image of

$$\overrightarrow{\text{chickens}} \otimes \overrightarrow{\text{cross}} \otimes \overrightarrow{\text{roads}} \quad (5)$$

under the map

$$\eta_N \otimes 1_S \otimes \eta_N : N \otimes (N \otimes S \otimes N) \otimes N \rightarrow S \quad (6)$$

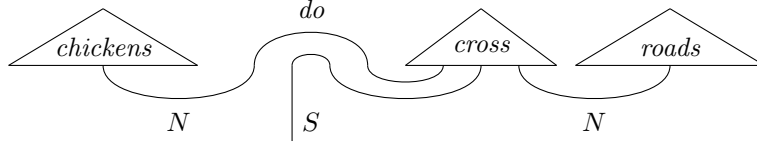
This nicely illustrates the general method. Our meaning category supplies the qualitative meanings of *chickens*, *cross*, and *roads*. Our grammar category then tells us how to stitch these together. This corresponds to ‘telling us where to put cups and caps.’ The essence of the method should be thought of as the diagram



where we think of the words as meaning vectors (5) and the wires as the map (6). In fact, rather than just using cups and identity wires, which are enough to account for the grammatical structure captured by pregroups, we can enrich the graphical language to directly account for meanings of functional words, such as relative pronouns.

## 2.6 Beyond standard categorial grammar

A first rather trivial example of a functional word is “does”, which can be accounted for by means of caps [11]:



Things become more interesting when we introduce, rather than just wires, the idea of a *multi-wire* (also called *spider* [8]) which can have more than two ends, or less:

$$\mu_N := \begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} \quad \iota_S := \begin{array}{c} \text{---} \\ \text{---} \end{array}$$

The way these behave is just like wires. The only thing that matters is: ‘either being connected by a multi-wire, or not’. As a consequence, multi-wires ‘fuse’ together:

$$\begin{array}{c} \text{---} \quad \text{---} \\ \text{---} \quad \text{---} \end{array} = \begin{array}{c} \text{---} \\ \text{---} \end{array}$$

These multi-wires can be defined in category-theoretic terms for any symmetric monoidal category, where they are called *commutative special dagger Frobenius structures* [10, 8].

**Example 7.** On each set  $X$  in the category **Rel** one can take the relations

$$X \times \dots \times X \rightarrow X \times \dots \times X :: \{((x, \dots, x), (x, \dots, x)) \mid x \in X\}$$

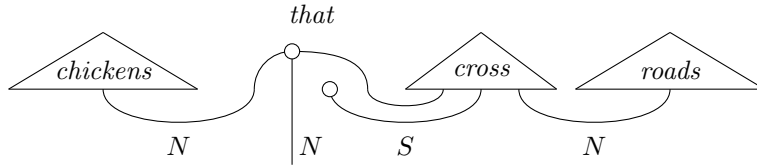
to be the multi-wires, and note in particular that these include identities, cups and caps.

**Example 8.** On each inner product space  $V$  in the category **FdVect** $_{\mathbb{R}}$ , given any orthonormal basis  $\{v_i\}_i$  for  $V$ , one can take the linear maps

$$V \otimes \dots \otimes V \rightarrow V \otimes \dots \otimes V :: v_{i_1} \otimes \dots \otimes v_{i_n} \mapsto \delta_{i_1 \dots i_n} v_{i_1} \otimes \dots \otimes v_{i_1}$$

to be the multi-wires, and again these include identities, cups and caps.

Using these multi-wires we can now express the meaning of relative pronouns [29, 30]:



Firstly, note that what we obtain is a noun rather than a sentence, and one that is closely related to ‘dead chicken’. Secondly, note that the use of the three-wire is mainly conjunction, conjoining ‘[is] chicken’ and ‘crosses road’. In this paper we will use them to directly express conjunctions. The one-wire gets rid of the sentence type.

### 3 Conceptual Spaces

*Conceptual spaces* are proposed in [12] as a framework for representing information at the conceptual level. Gärdenfors contrasts his theory with both a symbolic approach to concepts, and an associationist approach where concepts are represented as associations between different kinds of information elements. Instead, conceptual spaces are structures based on quality dimensions such as weight, height, hue and brightness. Concepts are roughly interpreted as convex subsets of a conceptual space. Conceptual spaces have an internal structure based on how quality dimensions interact with each other. A pair (or set) of dimensions is called *integral* if assignment of a value on one dimension requires assignment of a value on another dimension. For example, the dimensions of hue, saturation, and value in the HSV colour space are integral. Dimensions are called *separable* if values on one dimension can be assigned independently from the others. For example, hue and height are separable. The way in which this interaction is specified in Gärdenfors’s model is by using different distance metrics. Within a set of integral dimensions, distance is Euclidean, and between sets of integral dimensions, the city block metric is used.

Concept composition within conceptual spaces has been formalized in [26, 1, 22] for example. All these approaches focus on noun-noun composition, rather than utilising any more complex structure, and the way in which nouns compose often focuses on correlations between attributes in concepts. Since then, Gärdenfors has started to formalise verb spaces, adjectives, and other linguistic structures [13]. However, he has not provided a systematic method for how to utilise grammatical structures within conceptual spaces. In this sense, the category-theoretic approach to concept composition we describe below will introduce a more general approach to concept composition that can apply to varying grammatical types.

### 4 The Category of Convex Relations

In NLP applications, meanings are typically interpreted in categories of real vector spaces. For our intended cognitive application, we now introduce a category that emphasizes convex structure. The



familiar definition of convex set is a subset of a vector space which is closed under forming convex combinations. In this paper we consider a more general setting that includes convex subsets of vector spaces, but also allows us to consider some further, more discrete, examples.

We begin with some convenient notation. For a set  $X$  we write  $\sum_i p_i |x_i\rangle$  for a finite formal convex sum of elements of  $X$ , where  $p_i \in \mathbb{R}^{\geq 0}$  and  $\sum_i p_i = 1$ . We then write  $D(X)$  for the set of all such sums. Here we abuse the physicists ket notation to highlight that our sums are formal, following a convention introduced in [16]. Equivalently, these sums can be thought of as finite probability distributions on the elements of  $X$ .

A **convex algebra** is a set  $A$  with a function  $\alpha : D(A) \rightarrow A$  satisfying the following conditions:

$$\alpha(|a\rangle) = a \quad \text{and} \quad \alpha \left( \sum_{i,j} p_i q_{i,j} |a_{i,j}\rangle \right) = \alpha \left( \sum_i p_i | \alpha \left( \sum_j q_{i,j} |a_{i,j}\rangle \right) \rangle \right) \quad (7)$$

Informally,  $\alpha$  is a **mixing operation** that allows us to form convex combinations of elements, and the equations in (7) require the following good behaviour:

- Forming a convex combination of a single element  $a$  returns  $a$  as we would expect
- Iterating forming convex combinations interacts with flattening sums of sums in the way we would expect

We consider some examples of convex algebras.

**Example 9.** The closed real interval  $[0, 1]$  has an obvious convex algebra structure. Similarly, every real or complex vector space has a natural convex algebra structure using the underlying linear structure.

**Example 10 (Simplices).** For any set  $X$ , the formal convex sums of elements of  $X$  themselves form the **free convex algebra** on  $X$ , which can also be seen as a simplex with vertices the elements of  $X$ . Mixtures are formed as follows:

$$\sum_i p_i \left| \sum_j q_{i,j} |x_{i,j}\rangle \right\rangle \mapsto \sum_{i,j} p_i q_{i,j} |x_{i,j}\rangle$$

**Example 11.** The convex space of density matrices provides another example, with the convex structure given by the usual vector space structure on linear operators.

**Example 12.** For a set  $X$ , the functions of type  $X \rightarrow [0, 1]$  form a convex algebra pointwise, with mixing operation:

$$\sum_i p_i |f_i\rangle \mapsto (\lambda x. \sum_i p_i f_i(x))$$

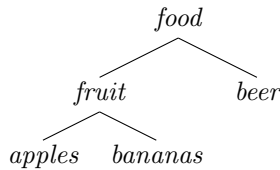
We can see this as a convex algebra of fuzzy sets.

**Example 13 (Semilattices).** As a slightly less straightforward example, every affine join semilattice (that is, one that has all finite *non-empty* joins) has a convex algebra structure given by:

$$\sum_i p_i |a_i\rangle = \bigvee_i \{a_i \mid p_i > 0\}$$

Notice that here the scalars  $p_i$  are discarded and play no active role. These “discrete” types of convex algebras allow us to consider objects such as the Boolean truth values.

**Example 14 (Trees).** Given a finite tree, perhaps describing some hierarchical structure, we can construct an affine semilattice in a natural way. For example, consider a limited universe of foods, consisting of bananas, apples, and beer. Given two members of the hierarchy, their join will be the lowest level of the hierarchy which is above them both. For instance, the join of *bananas* and *apples* would be *fruit*.



When  $\alpha$  can be understood from the context, we abbreviate our notation for convex combinations by writing:

$$\sum_i p_i a_i := \alpha(\sum_i p_i |a_i\rangle)$$

Using this convention, we define a **convex relation** of type  $(A, \alpha) \rightarrow (B, \beta)$  as a binary relation  $R : A \rightarrow B$  between the underlying sets that commutes with forming convex mixtures as follows:

$$(\forall i. R(a_i, b_i)) \Rightarrow R\left(\sum_i p_i a_i, \sum_i p_i b_i\right)$$

We note that identity relations are convex, and convex relations are closed under relational composition and converse.

**Example 15** (Homomorphisms). If  $(A, \alpha)$  and  $(B, \beta)$  are convex algebras, functions  $f : A \rightarrow B$  satisfying:

$$f\left(\sum_i p_i x_i\right) = \sum_i p_i f(x_i)$$

are convex relations. These functions are the **homomorphisms of convex algebras**. The identity function and constant functions are examples of homomorphisms of convex algebras.

The singleton set  $\{*\}$  has a unique convex algebra structure, denoted  $I$ . Convex relations of the form  $I \rightarrow (A, \alpha)$  correspond to **convex subsets**, that is, subsets of  $A$  closed under forming convex combinations.

**Definition 4.** We define the category **ConvexRel** as having convex algebras as objects and convex relations as morphisms, with composition and identities as for ordinary binary relations.

Given a pair of convex algebras  $(A, \alpha)$  and  $(B, \beta)$  we can form a new convex algebra on the cartesian product  $A \times B$ , denoted  $(A, \alpha) \otimes (B, \beta)$ , with mixing operation:

$$\sum_i p_i |(a_i, b_i)\rangle \mapsto \left(\sum_i p_i a_i, \sum_i p_i b_i\right)$$

This induces a symmetric monoidal structure on **ConvexRel**. In fact, the category **ConvexRel** has the necessary categorical structure for categorical compositional semantics:

**Theorem 1.** The category **ConvexRel** is a compact closed category. The symmetric monoidal structure is given by the unit and monoidal product outlined above. The caps for an object  $(A, \alpha)$  are given by:

$$\text{cap} : I \rightarrow (A, \alpha) \otimes (A, \alpha) :: \{(*, (a, a)) \mid a \in A\}$$

the cups by:

$$\text{cup} : (A, \alpha) \otimes (A, \alpha) \rightarrow I :: \{((a, a), *) \mid a \in A\}$$

and more generally, the multi-wires by:

$$\text{cap}_n : A \otimes \dots \otimes A \rightarrow A \otimes \dots \otimes A :: \{((a, \dots, a), (a, \dots, a)) \mid a \in A\}$$

**Remark 2.** As observed in remark 1, as **ConvexRel** is compact closed, its tensor cannot be a categorical product. For example, there are convex subsets of  $[0, 1] \times [0, 1]$  such as the diagonal:

$$\{(x, x) \mid x \in [0, 1]\}$$

that cannot be written as the cartesian product of two convex subsets of  $[0, 1]$ . This behaviour exhibits non-trivial *correlations* between the different components of the composite convex algebra.

**Remark 3.** We have given an elementary description of **ConvexRel**. More abstractly, it can be seen as the category of relations in the Eilenberg-Moore category of the finite distribution monad. Its compact closed structure then follows from general principles [5].

## 5 Adjective and Verb Concepts

We define a **conceptual space** to be an object of **ConvexRel**. In order to match the structure of the pregroup grammar, we require two distinct objects: a noun space  $N$  and a sentence space  $S$ .

The **noun space**  $N$  is given by a composite

$$N_{colour} \otimes N_{taste} \otimes \dots$$

describing different attributes such as colour and taste. A **noun** is then a convex subset of such a space. In our examples, we take our sentence space to be a convex algebra in which the individual points are events. Our general scheme can incorporate other sentence space structures, such choices are generally specific to the application under consideration. A **sentence** is then a convex subset of  $S$ .

We now describe some example noun and sentence spaces. We then show how these can be combined to form spaces describing adjectives and verbs. Once we have these types available, we show in section 6 how concepts interact within sentences.

### 5.1 Example: Food and Drink

We consider a conceptual space for food and drink as our running example. The space  $N$  is composed of the domains  $N_{colour}$ ,  $N_{taste}$ ,  $N_{texture}$ , so that

$$N = N_{colour} \otimes N_{taste} \otimes N_{texture}$$

The domain  $N_{colour}$  is the RGB colour domain, i.e. triples  $(R, G, B) \in [0, 1]^3$  with  $R, G, B$  standing for intensity of red, green, and blue light respectively.  $N_{taste}$  is defined as the simplex of combinations of four tastes: sweet, sour, bitter, and salt. We therefore have

$$N_{taste} = \{\vec{t} | \vec{t} = \sum_{i \in I} w_i \vec{t}_i\} \quad (8)$$

where  $I = \{sweet, sour, bitter, salt\}$ ,  $\vec{t}_i$  is the vector in some chosen basis of  $\mathbb{R}^4$  whose elements are all zero except for the  $i$ th element whose value is one, and  $\sum_i w_i = 1$ .  $N_{texture}$  is just the set  $[0, 1]$  ranging from completely liquid (0) to completely solid (1). We define a property  $p_{property}$  to be a convex subset of a domain, and specify the following examples (see figures 1 and 2):

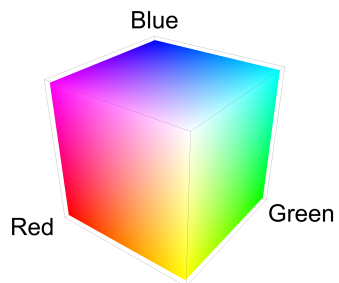
$$\begin{aligned} p_{yellow} &= \{(R, G, B) | (R \geq 0.7), (G \geq 0.7), (B \leq 0.5)\} \\ p_{green} &= \{(R, G, B) | (R \leq G), (B \leq G), (R \leq 0.7), (B \leq 0.7), (G \geq 0.3)\} \\ p_{sweet} &= \{\vec{t} | t_{sweet} \geq t_l \text{ for } l \neq sweet\} \end{aligned}$$

The properties  $p_{sour}$  and  $p_{bitter}$  are defined analogously.

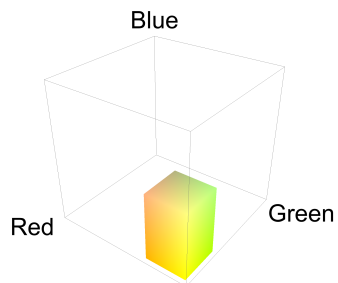
#### 5.1.1 Nouns

We define some nouns below. Properties in the colour domain are specified using sets of linear inequalities, and colours in the taste domain are specified using the convex hull of sets of points. We use  $Conv(A)$  to refer to the convex hull of a set  $A$ .

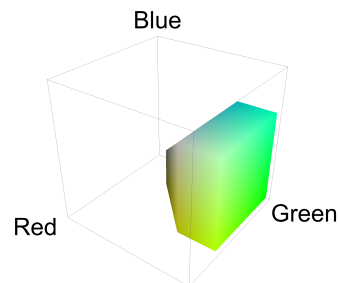
$$\begin{aligned} banana &= \{(R, G, B) | (0.9R \leq G \leq 1.5R), (R \geq 0.3), (B \leq 0.1)\} \\ &\quad \times Conv(\{t_{sweet}, 0.25t_{sweet} + 0.75t_{bitter}, 0.7t_{sweet} + 0.3t_{sour}\}) \times [0.2, 0.5] \\ apple &= \{(R, G, B) | R - 0.7 \leq G \leq R + 0.7, (G \geq 1 - R), (B \leq 0.1)\} \\ &\quad \times [0.5, 1] \times Conv(\{t_{sweet}, 0.75t_{sweet} + 0.25t_{bitter}, 0.3t_{sweet} + 0.7t_{sour}\}) \times [0.5, 0.8] \\ beer &= \{(R, G, B) | (0.5R \leq G \leq R), (G \leq 1.5 - 0.8R), (B \leq 0.1)\} \\ &\quad \times Conv(\{t_{bitter}, 0.7t_{sweet} + 0.3t_{bitter}, 0.6t_{sour} + 0.4t_{bitter}\}) \times [0, 0.01] \end{aligned}$$



(a) The RGB colour cube

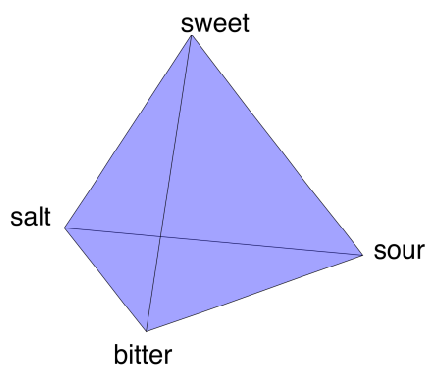


(b) Property  $p_{yellow}$

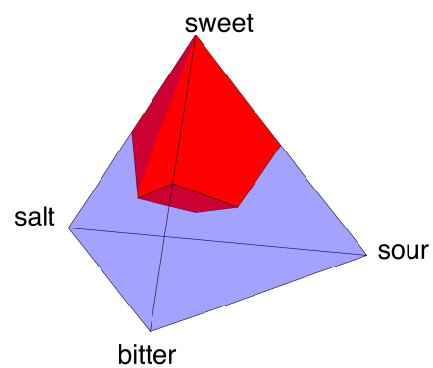


(c) Property  $p_{green}$

Figure 1: The RGB colour cube and properties  $p_{colour}$



(a) The taste tetrahedron



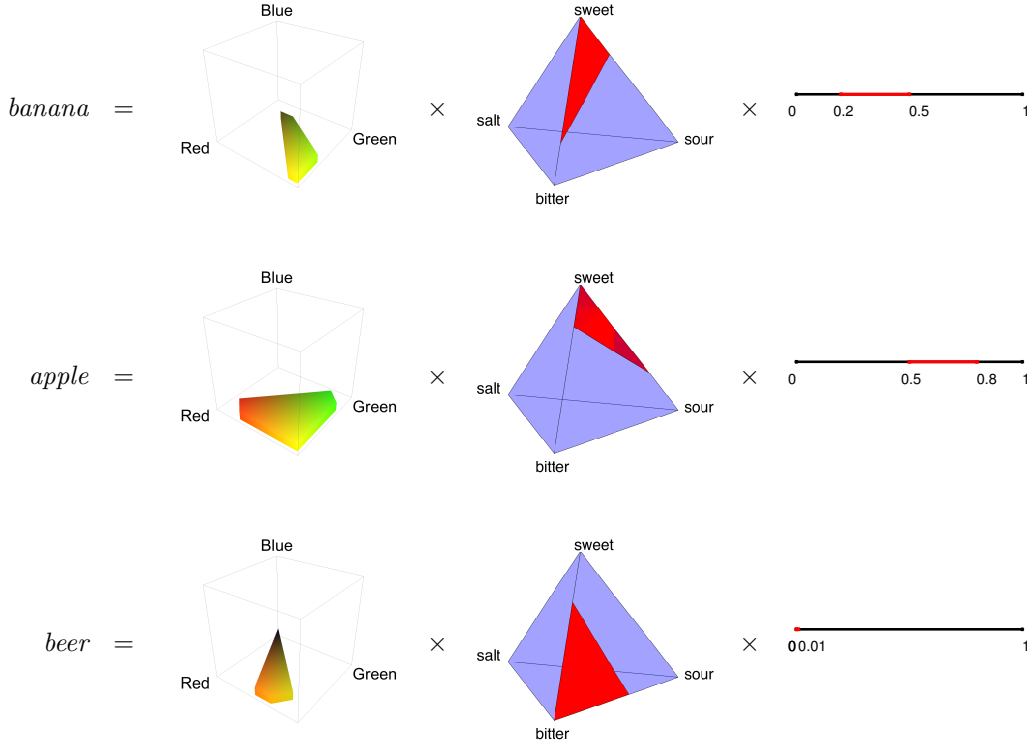
(b) The region corresponding to  $p_{sweet} = \{l | t_{sweet} \geq t_l \text{ for } l \neq \text{sweet}\}$

Figure 2: The taste space and the property  $p_{sweet}$

where  $t_i$  are as given in (8).

The subsets of points representing tastes are explained as follows using the case of banana as an example. Bananas are not at all salty, and therefore  $w_{salt}$  is set to 0. Bananas are sweet, and therefore the point  $t_{sweet}$  is chosen as an extremal point in the set of banana tastes. Bananas can also be somewhat but not totally bitter, and therefore the point  $0.25t_{sweet} + 0.75t_{bitter}$  is chosen as an extremal point. Similarly bananas can be a little sour, and therefore  $0.7t_{sweet} + 0.3t_{sour}$  is also chosen as an extremal point. Finally the convex hull of these points is formed giving a set of points corresponding to banana taste.

Pictorially, we have:



What is an appropriate choice of sentence space for describing food and drink? We need to describe the events associated with eating and drinking. We choose very simple structure where the events are either positive or negative, and surprising or unsurprising. We therefore use a sentence space of 2-tuples. The first element of the tuple states whether the sentence is positive (1) or negative (0) and the second states whether it is surprising (1) or unsurprising (0). The convex structure on this space is the convex algebra on a join semilattice induced by element-wise max, as in example 13. We therefore have four points in the space: positive, surprising (1, 1); positive, unsurprising (1, 0); negative, surprising (0, 1); and negative, unsurprising (0, 0). Sentence meanings are convex subsets of this space, so they could be singletons, or larger subsets such as  $negative = \{(0, 1), (0, 0)\}$ .

### 5.1.2 Adjectives

Recall that in a pregroup setting the adjective type is  $nn^l$ . In **ConvexRel**, the adjective therefore has type  $N \otimes N$ . Adjectives are convex relations on the noun space, so can be written as sets of ordered pairs. We give two examples,  $yellow_{adj}$  and  $soft_{adj}$ . The adjective  $yellow_{adj}$  has the simple form:

$$\{(\vec{x}, \vec{x}) | x_{colour} \in p_{yellow}\}$$

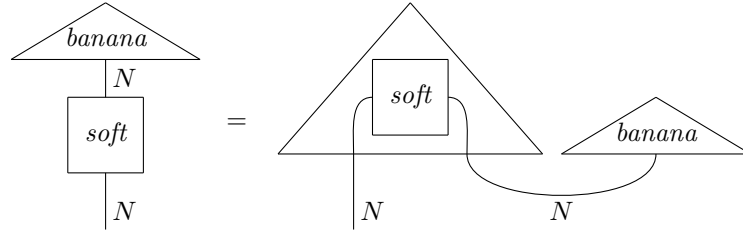
because it depends only on one area of the conceptual space.

An adjective such as ‘soft’ behaves differently to this. We cannot simply define soft as one area of the conceptual space, because whether or not something is soft depends what it was originally. Using relations, we can start to write down the right type of structure for the adjective, as long as the objects are sufficiently distinct. Restricting our universe just to bananas and apples, we can write  $soft_{adj}$  as

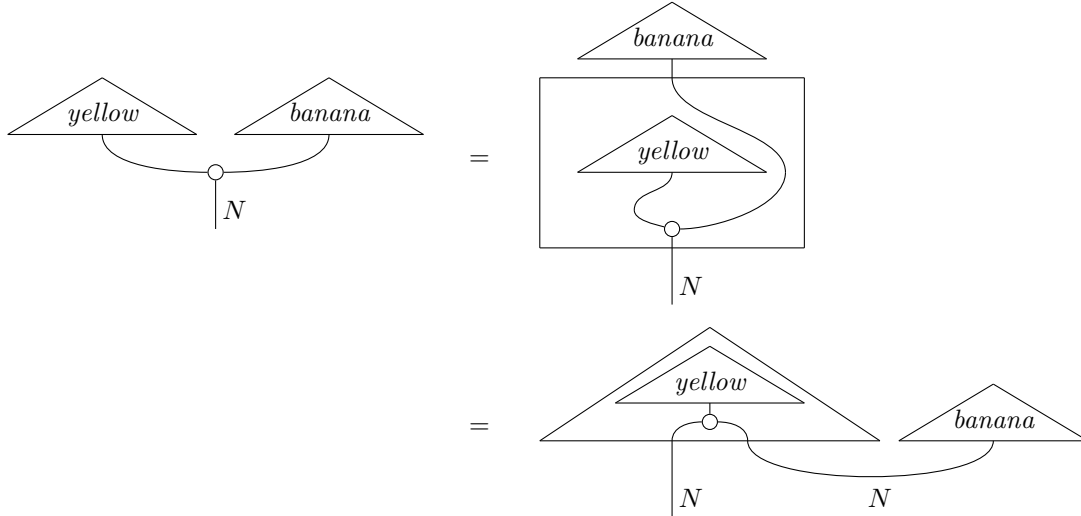
$$\{(\vec{x}, \vec{x}) | \vec{x} \in banana \text{ and } x_{texture} \leq 0.35 \text{ or } \vec{x} \in apples \text{ and } x_{texture} \leq 0.6\}$$

An analysis of the difficulties in dealing with adjectives set-theoretically, breaking them down into (roughly) three categories, is given in [17]. Under this view, both adjectives and nouns are viewed as one-place predicates, so that, for example  $red = \{x | x \text{ is red}\}$  and  $dog = \{x | x \text{ is a dog}\}$ . There are then three classes of adjective. For **intersective** adjectives, the meaning of  $adj \ noun$  is given by  $adj \cap \ noun$ . For **subjective** adjectives, the meaning of  $adj \ noun$  is a subset of  $\ noun$ . For **privative** adjectives, however,  $adj \ noun \not\subseteq \ noun$ .

Intersective adjectives are a simple modifier that can be thought of as the intersection between two concepts. We can make explicit the internal structure of these adjectives exploiting the multi-wires of theorem 1. For example, in the case of *yellow banana*, we take the intersection of *yellow* and *banana*. We then show how to understand *yellow* as an adjective. While the general case of adjectives is depicted as:



in the case of intersective adjectives the diagrams specialise to:



This shows us how the internal structure of an intersective adjective is derived directly from a noun.

### 5.1.3 Verbs

The pregroup type for a transitive verb is  $n^r sn^l$ , mapping to  $N \otimes S \otimes N$  in **ConvexRel**. To define the verb, we use concept names as shorthand, where these can easily be calculated. For example, since *green* is an intersective adjective, *green banana* is calculated by taking the intersection of *green* and *banana* by

combining the inequalities specifying the colour property, giving:

$$\begin{aligned} \text{green banana} = & \{(R, G, B) | (R \leq G \leq 1.5R), (G \geq B), (0.3 \leq R \leq 0.7), (G \geq 0.3), (B \leq 0.1)\} \\ & \times \text{Conv}(\{t_{\text{sweet}}, 0.25t_{\text{sweet}} + 0.75t_{\text{bitter}}, 0.7t_{\text{sweet}} + 0.3t_{\text{sour}}\}) \times [0.2, 0.5] \end{aligned}$$

Although a full specification of a verb would take in all the nouns it could possibly apply to, for expository purposes we restrict our nouns to just bananas and beer which do not overlap, due to the fact that they have different textures. We define the verb  $taste : I \rightarrow N \otimes S \otimes N$  as follows:

$$\begin{aligned} \text{taste} = & (\text{green banana} \times \{(0, 0)\} \times \text{bitter}) \cup (\text{green banana} \times \{(1, 1)\} \times \text{sweet}) \\ & \cup (\text{yellow banana} \times \{(1, 0)\} \times \text{sweet}) \cup (\text{beer} \times \{(0, 1)\} \times \text{sweet}) \cup (\text{beer} \times \{(1, 0)\} \times \text{bitter}) \end{aligned}$$

## 5.2 Example: Robot Navigation

We now present another example describing a simple formulation of robot navigation . We will describe our choices of noun space  $N$  and sentence space  $S$ , and show how to form nouns and verbs.

### 5.2.1 Nouns

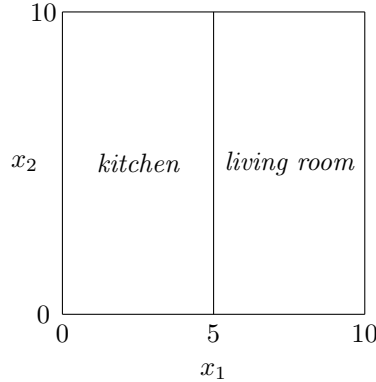
The types of nouns we wish to describe are objects, such as *armchair* and *ball*, the robots *Cathy* and *David*, and places such as *kitchen* and *living room*. For shorthand, we call these nouns  $a, b, c, d, k$ , and  $l$ . These are specified in the noun space  $N$  which is itself composed of a number of domains

$$N_{\text{location}} \otimes N_{\text{direction}} \otimes N_{\text{shape}} \otimes N_{\text{size}} \otimes N_{\text{colour}} \otimes \dots$$

We firstly consider the kitchen and living rooms as being defined by convex subsets of points in the domain  $N_{\text{location}}$ , defining properties in the location domain as:

$$\begin{aligned} p_{\text{kitchen location}} &= \{(x_1, x_2) | x_1 \in [0, 5], x_2 \in [0, 10]\} \\ p_{\text{living room location}} &= \{(x_1, x_2) | x_1 \in [5, 10], x_2 \in [0, 10]\} \end{aligned}$$

which can be depicted as follows:



Then the nouns *kitchen* and *living room* are given by these properties together with other sets of characteristics in the shape domain, size domain, and so on, which we won't specify here.

$$\begin{aligned} \text{kitchen} &= p_{\text{kitchen location}} \otimes p_{\text{kitchen shape}} \otimes p_{\text{kitchen size}} \otimes \dots \\ \text{living room} &= p_{\text{living room location}} \otimes p_{\text{living room shape}} \otimes p_{\text{living room size}} \otimes \dots \end{aligned}$$

Similarly, the other nouns are defined by combinations of properties in the noun space. For this example, we do not worry too much about what they are, but assume that they allow us to distinguish between the objects.

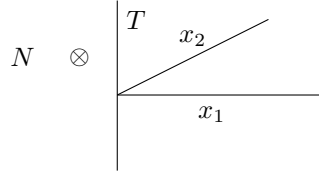
### 5.2.2 Verbs

In order to define some verbs, we need consider what a suitable sentence space should look like. We want to give sentences of the form:

The ball is in the living room  
Cathy moves to the kitchen

In these sentences, an object or agent is related to a path through time and space. Note that in the case of the verb ‘is in’, this path is in fact trivially just a point, however for ‘moves to’, the path actually is a path through time and space, and we will need to use subsets of the time and location domains to specify one single event. We therefore define the sentence space to be comprised of the noun space  $N$ , a time dimension  $T$ , and the location domain  $N_{location}$ :

$$S = N \otimes T \otimes N_{location}$$



The agent is represented by a point in the noun space  $N$ , and the path they take as described in the sentence is represented as a subset of the time and location domains.

In what follows, we think of 0 on the time dimension  $T$  as referring to ‘now’, with negative values of  $T$  referring to the past and positive values referring to the future.

As in the food example, transitive verbs are of the form  $N \otimes S \otimes N$ . This means that in this example, they are of the form

$$N \otimes N \otimes T \otimes N_{location} \otimes N,$$

and can be thought of as sets of ordered tuples of the form

$$(n_1, n_2, t, l, n_3),$$

where  $n_i$  stands for points in the noun space,  $t$  is a time,  $l$  is a location. We will consider the following verbs: *is\_in*, *moves\_to*. It could be argued that these are not transitive verbs, but intransitive verbs plus preposition. However, we can parse the combination as a transitive verb, since a preposition has type  $s^r sn^l$  and therefore the combination reduces to type of a transitive verb

$$(n^r s)(s^r sn^l) \leq n^r sn^l,$$

The verb *is\_in* can take any of the nouns  $a, b, c$ , or  $d$  as subject, and any of  $k, l$  as object. This verb refers to just one timepoint, i.e. now, or 0. The verb is as follows:

$$is\_in = \{(\vec{n}, \vec{n}, t_{now}, m_{location}, \vec{m}) | \vec{n} \in a \cup b \cup c \cup d, t_{now} = 0, \vec{m} \in k \cup l\} \quad (9)$$

The verb *moves\_to* refers to more than one point in time. We need to talk about an object moving from being at one location at a past time, to another location at time 0, or now. This movement should be continuous, since the objects we are talking about do not teleport from one point to another. We will also restrict the subject of the sentence to being one of the nouns  $a, b, c$ , or  $d$ , as we don’t want to talk about the kitchen and living rooms moving at this point. The object of the verb, however, can be any of the nouns, so we can say, for example, that ‘Cathy moves to the armchair’, or ‘The ball moves to Dave’ (presumably because Cathy kicked it). The most specific event that can be described in the space will track the exact path that an object takes through space and time. The meaning of a less specific sentence will be a convex subset of these trajectories. We now define the verb as follows:

$$moves\_to = \{(\vec{n}, \vec{n}, [t, 0], f([t, 0]), \vec{m}) \mid \vec{n} \in a \cup b \cup c \cup d, t < 0, f \text{ continuous}, f(t) \in \vec{n}_{location}, f(0) \in \vec{m}_{location}\} \quad (10)$$



The constraints on  $t$  ensure that the movement happened in the past, and the constraints on  $f$  ensure that the movement is from the location of the subject to the location of the object of the verb. These definitions are a little complex, but we will see how they work in interaction. Note that the location nouns *kitchen* and *living room* might seem to be of a different type to the object and agent nouns *armchair*, *ball*, *Cathy* and *David*. For example, we have specified that *kitchen* and *living room* do not move around. In future research we will be extending to a richer type system which can take account of these sorts of differences, and which will in fact be closer to that proposed by Gärdenfors in [13].

## 6 Concepts in Interaction

We have given descriptions of how to form the different word types within our model of categorical conceptual spaces. In this section we show how to apply the type reductions of the pregroup grammar within the conceptual spaces formalism.

### 6.1 Sentences in the Food Space

The application of  $yellow_{adj}$  to *banana* works as follows.

$$\begin{aligned}
yellow\ banana &= (1_N \times \epsilon_N)(yellow_{adj} \times banana) \\
&= (1_N \times \epsilon_N)\{(\vec{x}, \vec{x}) | x_{colour} \in yellow\} \\
&\quad \times \{(R, G, B) | (0.9R \leq G \leq 1.5R), (R \geq 0.3), (B \leq 0.1)\} \\
&\quad \times Conv(\{t_{sweet}, 0.25t_{sweet} + 0.75t_{bitter}, 0.7t_{sweet} + 0.3t_{sour}\}) \times [0.2, 0.5] \\
&= \{(R, G, B) | (0.9R \leq G \leq 1.5R), (R \geq 0.7), (G \geq 0.7), (B \leq 0.1)\} \\
&\quad \times Conv(\{t_{sweet}, 0.25t_{sweet} + 0.75t_{bitter}, 0.7t_{sweet} + 0.3t_{sour}\}) \times [0.2, 0.5]
\end{aligned}$$

Notice, in the last line, how the colour property has altered. This alteration restricts to yellow hues. This assumes that we can tell bananas and apples apart by shape, colour and so on. Then the same calculation gives us

$$\begin{aligned}
soft\ apple &= \{(R, G, B) | R - 0.7 \leq G \leq R + 0.7, (G \geq 1 - R), (B \leq 0.1)\} \\
&\quad \times Conv(\{t_{sweet}, 0.75t_{sweet} + 0.25t_{bitter}, 0.3t_{sweet} + 0.7t_{sour}\}) \times [0.4, 0.6]
\end{aligned}$$

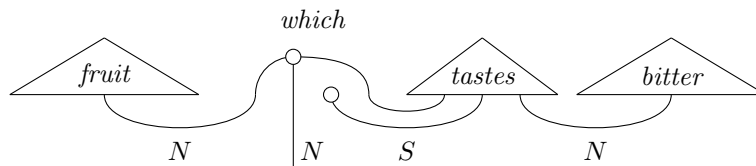
Using the definition of *taste* that we gave, we find that although sweet bananas are good:

$$\begin{aligned}
bananas\ taste\ sweet &= (\epsilon_N \times 1_S \times \epsilon_N)(bananas \times taste \times sweet) \\
&= (\epsilon_N \times 1_S)(banana \times (green\ banana \times \{(1, 1)\} \cup yellow\ banana \times \{(1, 0)\})) \\
&= \{(1, 1), (1, 0)\} \\
&= positive
\end{aligned}$$

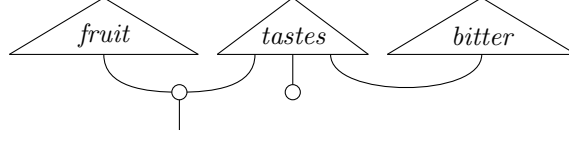
sweet beer is not so desirable:

$$\begin{aligned}
beer\ tastes\ sweet &= (\epsilon_N \times 1_S \times \epsilon_N)(beer \times taste \times sweet) \\
&= \{(0, 1)\} \\
&= negative\ and\ surprising
\end{aligned}$$

**Relative Pronouns** The compositional semantics we use can also deal with relative pronouns, described in detail in [19]. Relative pronouns are words such as ‘which’. For example, we can form the noun phrase *Fruit which tastes bitter*. This has the following structure:



which simplifies to:



In our example, we find that *Fruit which tastes bitter = green banana*:

$$\begin{aligned}
\text{Fruit which tastes bitter} &= (\mu_N \times \iota_S \times \epsilon_N)(\text{Conv}(\text{bananas} \cup \text{apples}) \times \text{taste} \times \text{bitter}) \\
&= (\mu_N \times \iota_S)(\text{Conv}(\text{bananas} \cup \text{apples}) \times (\text{green banana} \times \{(0,0)\})) \\
&= \mu_N(\text{Conv}(\text{bananas} \cup \text{apples}) \times (\text{green banana})) \\
&= \text{green banana}
\end{aligned}$$

where  $\mu_N$  is the converse of the Frobenius copy map on  $N$  and  $\iota_S$  is the delete map on  $S$  from theorem 1.

## 6.2 Sentences about Robot Movement

In this section we describe how to compute the meaning of sentences about robot movement. Our first example is the sentence ‘Cathy moves to the living room’. In order to compute the meaning of this sentence, we assume that Cathy has a location.

$$\begin{aligned}
&\text{Cathy moves to the living room} \\
&= (\epsilon_N \otimes 1_s \otimes \epsilon_N)(C \otimes \text{moves.to} \otimes L) \\
&= (\epsilon_N \otimes 1_s \otimes \epsilon_N)(C \otimes \{(\vec{n}, \vec{n}, [t, 0], f([t, 0])) | f(0) \in L_{\text{location}}\}) \\
&= \{C, [t, 0], f([t, 0]) | f(t) \in C_{\text{location}}, f(0) \in L_{\text{location}}\}
\end{aligned} \tag{11}$$

In line (11) further constraints apply to  $t$  and  $f$  as described in equation (10). This calculation gives us a set of continuous line segments starting from Cathy’s location at time  $t$  and ending in the living room at time 0.

We now need to check that this set of line segments is convex. We assume that Cathy can take any possible location, and her other attributes remain static. This means that the set of possible Cathys is convex. The set of time segments  $[t, 0]$  such that  $t < 0$  forms a convex set. Consider two such time segments. We define a convex mixture of these segments pointwise:

$$p[t_1, 0] + (1 - p)[t_2, 0] = [pt_1 + (1 - p)t_2, p0 + (1 - p)0] = [pt_1 + (1 - p)t_2, 0]$$

which clearly satisfies the condition that the start point is in the past and the end point is now.

We then consider the convex mixture of two sets of locations  $f_1([t_1, 0])$  and  $f_2([t_2, 0])$ . In order to carry this out, we first of all transform the intervals  $[t_1, 0]$  and  $[t_2, 0]$  to  $[-1, 0]$  by dividing through by  $-t_i$ , renaming the rescaled functions  $f'_i$ . We then form a convex combination:

$$pf'_1 + (1 - p)f'_2 : [0, 1] \rightarrow N_{\text{location}}$$

pointwise by taking:

$$(pf'_1 + (1 - p)f'_2)(\tau) = pf'_1(\tau) + (1 - p)f'_2(\tau) = pf_1((-t_1)\tau) + (1 - p)f_2((-t_2)\tau)$$

Since both  $f_1$  and  $f_2$  are continuous in  $T$ , the result will be continuous in  $T$ .

The constraints on these sets are that  $f_1(t_1)$  and  $f_2(t_2)$  are in  $C_{\text{location}}$  and that  $f_1(0)$  and  $f_2(0)$  are in  $L_{\text{location}}$ , and we need that their convex combinations are also in these respective locations. We know that  $C_{\text{location}}$  and  $L_{\text{location}}$  are convex, meaning that

$$pf_1(t_1) + (1 - p)f_2(t_2) \in C_{\text{location}} \quad \text{and} \quad pf_1(0) + (1 - p)f_2(0) \in L_{\text{location}}$$

as required.

In this section, we have shown how we can use the interaction of words, represented by convex sets in conceptual spaces, to map sentence meanings down to a convex set in a conceptual space for sentences.

## 7 Conclusion

We have applied the categorical compositional scheme to cognition and conceptual spaces. In order to do this we introduced a new model for categorical compositional semantics, the category **ConvexRel** of convex algebras and binary relations respecting convex structure. We consider this model as a proof of concept that we can describe convex structures within our framework. Conceptual spaces are often considered to have further mathematical structure such as distance measures and notions of convergence or fixed points. It is also possible to vary the notion of convexity under consideration, for example by considering a binary betweenness relation on a space as primitive, rather than a mixing operation. Identifying a good setting for rich conceptual spaces models, and incorporating those structures into a compositional framework is a direction for further work.

## Acknowledgements

This work was partially funded by AFSOR grant “Algorithmic and Logical Aspects when Composing Meanings”, the FQXi grant “Categorical Compositional Physics”, and EPSRC PhD scholarships.

## References

- [1] B. Adams & M. Raubal (2009): *A metric conceptual space algebra*. In: *Spatial information theory*, Springer, pp. 51–68, doi:10.1007/978-3-642-03832-7\_4.
- [2] D. Bankova, B. Coecke, M. Lewis & D. Marsden (2015): *Graded Entailment for Compositional Distributional Semantics*. arXiv:1601.04908. Submitted.
- [3] J. Bolt, B. Coecke, F. Genovese, M. Lewis, D. Marsden & R. Piedeleu (2016): *Interacting Conceptual Spaces*. In D. Kartsaklis, M. Lewis & L. Rimell, editors: *Proceedings of the 2016 Workshop on Semantic Spaces at the Intersection of NLP, Physics and Cognitive Science, SLPCS@QPL 2016, Glasgow, Scotland, 11th June 2016., EPTCS 221*, pp. 11–19, doi:10.4204/EPTCS.221.2. Available at <http://dx.doi.org/10.4204/EPTCS.221.2>.
- [4] J.A. Bullinaria & J.P. Levy (2007): *Extracting semantic representations from word co-occurrence statistics: A computational study*. *Behavior research methods* 39(3), pp. 510–526.
- [5] A. Carboni & R.F.C. Walters (1987): *Cartesian bicategories I*. *Journal of pure and applied algebra* 49(1), pp. 11–32, doi:10.1016/0022-4049(87)90121-6.
- [6] B. Coecke (2013): *An alternative Gospel of structure: order, composition, processes*. In C. Heunen, M. Sadrzadeh & E. Grefenstette, editors: *Quantum Physics and Linguistics. A Compositional, Diagrammatic Discourse*, pp. 1–22.
- [7] B. Coecke, E. Grefenstette & M. Sadrzadeh (2013): *Lambek vs. Lambek: Functorial vector space semantics and string diagrams for Lambek calculus*. *Annals of Pure and Applied Logic* 164(11), pp. 1079–1100, doi:10.1016/j.apal.2013.05.009.
- [8] B. Coecke & A. Kissinger (2016): *Picturing Quantum Processes. A First Course in Quantum Theory and Diagrammatic Reasoning*. Cambridge University Press.
- [9] B. Coecke & E.O. Paquette (2011): *Categories for the practising physicist*. In: *New Structures for Physics*, Springer, pp. 173–286, doi:10.1007/978-3-642-12821-9.
- [10] B. Coecke, D. Pavlović & J. Vicary (2013): *A new description of orthogonal bases*. *Mathematical Structures in Computer Science, to appear* 23, pp. 555–567. arXiv:quant-ph/0810.1037.
- [11] B. Coecke, M. Sadrzadeh & S. Clark (2010): *Mathematical Foundations for a Compositional Distributional Model of Meaning*. *Linguistic Analysis* 36, pp. 345–384.

- [12] P. Gärdenfors (2004): *Conceptual spaces: The geometry of thought*. The MIT Press.
- [13] P. Gärdenfors (2014): *The geometry of meaning: Semantics based on conceptual spaces*. MIT Press.
- [14] E. Grefenstette & M. Sadrzadeh (2011): *Experimental Support for a Categorical Compositional Distributional Model of Meaning*. In: *The 2014 Conference on Empirical Methods on Natural Language Processing*, pp. 1394–1404. ArXiv:1106.4058.
- [15] J. Hampton (1987): *Inheritance of attributes in natural concept conjunctions*. *Memory & Cognition* 15(1), pp. 55–71.
- [16] B. Jacobs (2011): *Coalgebraic walks, in quantum and Turing computation*. In: *Foundations of Software Science and Computational Structures*, Springer, pp. 12–26, doi:10.1007/978-3-642-19805-2\_2.
- [17] H. Kamp & B. Partee (1995): *Prototype theory and compositionality*. *Cognition* 57(2), pp. 129–191, doi:10.1016/0010-0277(94)00659-9.
- [18] D. Kartsaklis & M. Sadrzadeh (2013): *Prior disambiguation of word tensors for constructing Sentence vectors*. In: *The 2013 Conference on Empirical Methods on Natural Language Processing*, ACL, pp. 1590–1601.
- [19] D. Kartsaklis, M. Sadrzadeh, S. Pulman & B. Coecke (2013): *Reasoning about meaning in natural language with compact closed categories and Frobenius algebras*. In J. Chubb, A. Eskandarian & V. Harizanov, editors: *Logic and Algebraic Structures in Quantum Computing*, Cambridge University Press (CUP), pp. 199–222, doi:10.1017/cbo9781139519687.011.
- [20] J. Lambek (1958): *The mathematics of sentence structure*. *American Mathematics Monthly* 65.
- [21] J. Lambek (1999): *Type grammar revisited*. In: *Logical aspects of computational linguistics*, Springer, pp. 1–27, doi:10.1007/3-540-48975-4\_1.
- [22] M. Lewis & J. Lawry (2016): *Hierarchical conceptual spaces for concept combination*. *Artificial Intelligence* 237, pp. 204 – 227, doi:10.1016/j.artint.2016.04.008. Available at <http://www.sciencedirect.com/science/article/pii/S0004370216300492>.
- [23] K. Lund & C. Burgess (1996): *Producing high-dimensional semantic spaces from lexical co-occurrence*. *Behavior Research Methods, Instruments, & Computers* 28(2), pp. 203–208, doi:10.3758/BF03204766.
- [24] S. Mac Lane (1971): *Categories for the Working Mathematician*. Springer.
- [25] R. Piedeleu, D. Kartsaklis, B. Coecke & M. Sadrzadeh (2015): *Open System Categorical Quantum Semantics in Natural Language Processing*. In Lawrence S. Moss & Pawel Sobocinski, editors: *6th Conference on Algebra and Coalgebra in Computer Science, CALCO 2015, LIPIcs* 35, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, pp. 270–289, doi:10.4230/LIPIcs.CALCO.2015.270.
- [26] J.T. Rickard, J. Aisbett & G. Gibbon (2007): *Reformulation of the theory of conceptual spaces*. *Information Sciences* 177(21), pp. 4539–4565, doi:10.1016/j.ins.2007.05.023.
- [27] Eleanor Rosch & Carolyn B. Mervis (1975): *Family resemblances: Studies in the internal structure of categories*. *Cognitive psychology* 7(4), pp. 573–605.
- [28] M. Sadrzadeh (2007): *High Level Quantum Structures in Linguistics and Multi Agent Systems*. In: *AAI Spring Symposium: Quantum Interaction*, pp. 9–16.
- [29] M. Sadrzadeh, S. Clark & B. Coecke (2013): *The Frobenius anatomy of word meanings I: subject and object relative pronouns*. *Journal of Logic and Computation* 23, pp. 1293–1317. ArXiv:1404.5278.

- [30] M. Sadrzadeh, S. Clark & B. Coecke (2014): *The Frobenius anatomy of word meanings II: possessive relative pronouns*. *Journal of Logic and Computation*, p. exu027.
- [31] A. Tversky (1977): *Features of similarity*. *Psychological Review* 84(4), pp. 327–352.